

## Two controlled trials of supplementary feeding of British school children in the 1920s

Jon Pollock

*J R Soc Med* 2006;99:323–327

*Additional material for this article is available from the James Lind Library website [www.jameslindlibrary.org] where this paper was previously published.*

In 1971, Margaret Thatcher, then Minister of State for Education in the Edward Heath government, introduced the *Education (Milk) Act*. This abolished universal free school milk, which had been an important plank in the 'nutrition for education' policies introduced from the early 1900s. The ecological relationship between low rates of breast-feeding and high infant mortality, and the generally poor physical condition of young people living in poverty, had led to the *Education (Provision of Meals) Act* in 1906. This released funds to provide nutrition for children unable to take full advantage of educational opportunities through lack of food. In fact, the educational impact of school meals and milk was never properly evaluated.

Milk is the nurturing medium for infants and is symbolic of motherhood, growth and reproduction as the 'complete food'. Although they have wide nutritive potential, however, milk supplements can also have negative impacts, both as substitutes for higher quality solid foods,<sup>1</sup> and because of poor sterilization. Furthermore, the dairy industry may also have influenced policy.<sup>2,3</sup> Although cow's milk was relatively inexpensive and logistically easy to distribute, it was realized after the First World War that its nutritional impact as a supplement should be properly assessed if milk distribution was to become part of a nationwide policy initiative.

### MCCOLLUM'S STUDY IN THE NEGRO ORPHANS' HOME IN BALTIMORE, USA

In the United States, Elmer McCollum, a nutritionist at the Johns Hopkins School of Hygiene and Public Health who had previously played a key role in discovering Vitamins A, B and D, undertook a controlled trial of supplementary milk (2 pints of reconstituted dried milk daily) in the Negro Orphans' Home in Baltimore.<sup>4</sup> Many of the children who participated in the study had '... signs of tuberculosis [and] rickets', and were described as 'severely malnourished'

(only children with signs of syphilis were excluded from the study). Eighty-four children aged 4 to 10 were divided into 42 in an intervention arm and 42 controls. McCollum does not report the method used to allocate the 84 children to one or other of the two groups, but he states that '... every effort was made ... so that any child in one group was comparable in age, size and condition to a child in the other group'. The children had a basic diet of cereals and a soup made from root and tuber vegetables (a maximum of 1500 calories/child/day), which McCollum acknowledged was probably insufficient for proper growth. The study appears to have been successful, but it has been criticized for being confounded by the 'Hawthorne' effect because a major change in the basic diet was introduced as a result of the study taking place. In fact, the presentation of the results is very confusing: no proper statistical analysis is presented, and the two comparison groups seemingly differed substantially in their basic diet.

### CORRY MANN'S STUDY AT THE BARNARDO'S 'COLONY OF BOYS' AT WOODFORD, ESSEX, ENGLAND

In the UK, small studies such as that of Auden<sup>5</sup> were also attempting to assess the physical effects of supplementary milk rations. Eventually, Dr AWJ MacFadden, a Senior Medical Officer in the Ministry of Health, proposed a study on the 'effects produced in the growth and nutrition of *underdeveloped* (my italics) children by the addition of milk to a standard diet'. This proposal was approved in May 1921 at a meeting of the Medical Research Council's Accessory Food Factors Committee. Harold Corry Mann, who was employed by the Medical Research Council, was charged with implementing the study, which was eventually published in 1926 under the title of *Diets for Boys During the School Age*.<sup>6</sup> The report was at pains to demonstrate that the basic diet provided was of '... adequate physiological value ... (with) ... additional calories for the *minimum* (my italics) requirements of growth and activity ...'. Perhaps the slightly contradictory wording was designed to mitigate any accusations that the institutions involved were not providing a satisfactory basic diet in the first place.

The design of the study, as reported, is somewhat confusing. Basically, it consisted of a multi-arm, controlled, non-randomized trial of whole milk, milk components, and

other supplements, and involved groups of boys aged between 6 and 12 years of age, for periods up to 3 years. The participants lived in Barnardo's 'colony of boys' at Woodford, Essex, 11 miles from London, which provided full-time care for those in need. The colony was a self-contained 'model village', with its own school, hospital, farm, swimming pool, bakery and disinfecting plant. The 500 to 600 resident boys lived in houses, each of which had about 30 residents. They ate their meals communally in a large dining hall, although, when outbreaks of infectious disease occurred, houses could be isolated and managed individually by house-matrons using their own kitchens. Boys were selected for the experiment according to age (7 to 11) and weight (45 to 65lb), but were excluded if they were 'of colour or of foreign, Latin, Scandinavian, or Hebrew type', or if they had obvious disease. The boys were 'rated' on height and weight according to a scale of 'A1 to B3'.

This attention to racial detail, presumably on the grounds of minimizing genetic variation, contrasts with the lack of detail about how boys were allocated to different houses, and thereby to different diets. The report states that '... as far as possible an equal number of the same age and rating were assigned to each of the three houses. . .', but it is confusing about the changing numbers of boys per house (departures and intakes, and the likely bias this could introduce), the duration of supplementation, and changing age and weight criteria.

The study began with a 4-month period of baseline data collection on diet and physical measurements and activity in the three houses in which selected boys were lodged for this phase of the experiment. Boys occupying seven of the eight houses were then allocated one of five different supplements, in addition to the basic diet delivered in the communal dining hall. Those occupying one house received no supplement and those in two of the other houses received the same supplement. No information was provided on how houses were allocated to supplements. The boys of each house presumably spent more time with one another as well as occupying the same dwelling, and (it appears) they actually ate together at tables within the communal dining room. However, there is no reason to suspect that they were 'clustered' in the sense of the basic diet received.

The nutritional composition of the basic diet was assessed at the time of consumption by selecting 'three or four plates from each table of 30 boys' to provide a range of itemized weights of each food type. It is unclear how often this was done and whether it continued during the supplementation period and for all supplementation groups. The supplements themselves were analysed for nutrients but the report provides little detail on the sampling frames and frequencies.

Every day, each boy received a supplement consisting of 1 pint of fresh but pasteurized cow's milk, 3 oz of sugar, 1¾ oz butter, 1¾ oz vegetable margarine, ¾ oz casein, and, for those in two of the houses, ½ to ¾ oz watercress. The caloric value of the milk, butter, sugar and margarine supplements was almost identical at between 350 to 388 calories. The weights of the boys were measured every 2 weeks; their heights every 3 months.

The results of the experiment indicated that, compared with other groups, boys receiving whole milk supplements and butter supplements experienced greater gains in weight and height. These differences were not directly analysed statistically. Instead, statistical analysis compared the differences between observed and expected changes in weight and height, based on the Null hypothesis of no difference. Expected values, stratified by starting weight band and by season (summer/winter), were derived from regression equations developed from a study of 9146 boys aged 4.5 to 14 years in rural schools.

Despite relatively small numbers, the overall pattern of results indicated that both milk and butter supplementation were associated with similar, statistically significant increments in weight gain (about 11 oz) over expected values per 6-month period of exposure, with additional gains in height of about 0.38 inches in the milk supplemented group and 0.17 inches in the butter supplemented group. No discussion of these results is offered. Nor is reference made to the statistically significant *negative* observed minus expected gain in weight of about 7.6 oz in the basic diet group.

The likely origin of the regression data in cross-sectional rather than longitudinal studies, and the inflation of sample size through the repeated measures using the same children in 6-month blocks, would not inspire confidence in these results today. But what should we make of these results, given that they were instrumental in the implementation of the Milk in Schools Scheme a few years after they were published? The strengths of the study undoubtedly lie in its baseline data collection, the duration of the supplementation period, the frequency of outcome measurement, and the great attention to detail in presenting the itemized diets, even on a daily basis. Its weaknesses are apparent in inadequate descriptions of the selection and exclusion criteria, the means of allocation to supplemented groups, and the poor and confusing explanation of changing criteria and follow-up routines. Did healthy, well-built children find placements away from the colony and the experiments earlier than others, and would this have occurred independently of their supplementation group? The report tells us nothing about individual food intake measures in a social situation where, unblinded, children from known houses presumably queue up together in line for a plate of food provided by serving staff who were also well aware of

the supplementation regimens. This might work in favour of the unsupplemented children, but hardly explains the worrying deficit in expected weight gain in this group. This lends credibility to a concern that the basic diet was suboptimal, and that supplemented children were 'catching up' with expectation.

Criticisms of Corry Mann's research were published at the time, but these mostly related to concerns over the duration and costs of the experiment which, in 1 year, accounted for 3% of the Medical Research Council's annual research budget (Atkins, personal communication). There appears to have been little contemporaneous scientific criticism. Petty, a more recent critic,<sup>1</sup> claims that Corry Mann's research is seriously flawed because the observed effects represent only the results of 'catch-up' growth in the milk supplemented group, which had a greater proportion of stunted boys; and she suggests that extra growth was related to calorific value rather than type of supplement. However, the lack of a statistically significant difference found between milk and basic diet groups in the 'non-stunted' category of boys that she constructed using Corry Mann's data, may reflect small sample sizes rather than small effect sizes. It is true that Corry Mann's stratification by weight rather than by age may have obscured differences in growth impairment between the comparison groups. However, further analysis of the data presented in the paper does not reveal any statistically significant differences at baseline between boys receiving the basic diet and those receiving milk supplements in either height-for-weight ( $t=1.11$ ,  $P=0.267$ ,  $df=100$ ) or height-for-age ( $t=1.59$ ,  $P=0.115$ ,  $df=100$ ). Furthermore, in the only age bands with reasonable numbers for age-stratified analysis, weight and height gain were superior in the group receiving milk supplements compared with those receiving only the basic diet (9 year olds: weight:  $t=6.70$ ,  $P<0.001$ ,  $df=26$ ; height:  $t=3.91$ ,  $P=0.002$ ,  $df=11$ ; 10 year olds: weight:  $t=4.68$ ,  $P<0.001$ ,  $df=15$ ; height:  $t=5.45$ ,  $P<0.001$ ,  $df=15$ ; and 11 year olds: weight:  $t=5.71$ ,  $P<0.001$ ,  $df=16$ ; height:  $t=6.23$ ,  $P<0.001$ ,  $df=11$ ).

I chose these three age groups primarily because they provided sufficient numbers of children for analysis, but I had another reason: the weight-dependent inclusion criteria introduced an age discrepancy in stature which is likely to have operated to the disadvantage of the older children. Thus, using the relatively modern standard of the National Center for Health Statistics Growth Curves, a 7-year-old boy meeting the Corry Mann weight criterion (45 to 65 lb) would today lie between the 10th and 90th centile; an 11 year old would lie between the 3rd and the 15th centile. Irrespective of the standard employed, however, it is very likely that at least some nutritionally deprived older boys were included in the experiment, making an age-stratified analysis advisable. It seems quite possible, as Petty<sup>1</sup> also

concluded, that the real result of the trial was to show the effect of providing extra wholesome food containing fats, carbohydrates and protein to supplement a Barnardo's diet that was probably inadequate in the first place.

After making a qualitative comparison of the institutionalised boys in his Barnardo's sample with other boys living in the area, Corry Mann went on to infer the generalizability of his results outside the context of the home. He made no reference to the importance of vitamins and other components of milk, simply presenting the results of milk versus sugar versus casein versus butter and margarine (and watercress), and leaving the implications to be drawn by others. Subsequent decisions to invest in free milk provision for school age children, rather than, for example, reviewing the dietary adequacy of institutionalized children, might seem strange today. But, as Atkins (personal communication) and others have demonstrated, policy based solely on flawed research findings has a long history.

#### STUDIES BY ORR, LEIGHTON AND CLARK IN SCOTLAND AND BELFAST

The challenge of generalizing from the institution-based trials done by McCollum and Corry Mann to the wider community, where dietary input could not be carefully controlled, led to a large study based in day schools in Scotland and Belfast.<sup>7-9</sup> Starting late in 1926, 1425 children 'living in the ordinary conditions of Scottish working class homes', aged 5 to 13 years and distributed in six Scottish towns and in Belfast, were assigned to one of four dietary supplements for two periods, each of 7 months, followed by 5 months 'rest', over 2 years. It was intended that children in the four groups would receive: (i)  $\frac{3}{4}$  to  $1\frac{1}{4}$  pints (according to age) whole milk (3.85% fat); (ii) the same amounts of 'separated' (mechanically skimmed) milk (0.33% fat); (iii) a biscuit of the same caloric value as the skimmed milk; or (iv) nothing. Heights and weights were recorded at the end of each year.

The implication from the report was that the intervention arms were allocated by school class, as each group was numbered '... from 25 to 50 according to the size of the classes in that school'.<sup>9</sup> However, the report tells us only that '... at each place 4 groups of children were selected and each group treated differently'.<sup>8</sup> Although it appears that each of the four study arms was represented in each school, which would probably control for social and economic variation between arms, it seems unlikely that these were of the same age group. In any case, the analysis was by age group across schools and did not take variation between schools into account.

Comparison of the increases in height and weight between children who received at least 75% of the milk supplements (1282 in year 1 and 1157 in year 2) with

children in the biscuit and control arms indicated an advantage, sustained over one and two years, among the former. Little difference was detected between the effects of the different milk types, or between those associated with biscuit and controls. This led the researchers to conclude that milk supplementation at this level was of 'wide public health significance'; that skimmed milk was generally equivalent to whole milk and superior to biscuit; and that the advantage was therefore not attributable to differential fat or calorie intake. The trial is of additional interest in that it claimed to demonstrate reversibility of the effects in a cross-over experiment during the second year, in which three groups of children were assigned to different arms of the trial from those to which they had been assigned initially. However, the numbers involved in this part of the trial must have been small and no statistical tests were applied.

It is unlikely that the possible role of selection bias was seen as undermining the credibility and importance of this trial, which was well received by doctors. However, the Scottish Department of Health commissioned a further, larger study early in 1930, on the grounds that the '... improvement in nutrition of the children who received the additional ration of milk was due ... in some measure to improved home conditions ... which might follow from close surveillance ...' of only the intervention group,<sup>9</sup> a recognition (albeit on uncertain foundations) that the Hawthorne effect may have been operating. Perhaps politicians of the day feared the economic impact of universal dietary supplementation in schools and needed final 'proof' of effect. Under the *Education (Scotland) Act* of 1930 local authorities were granted the right to provide additional milk rations where appropriate and this could have affected 800 000 children of school age. The objective, as defined by the Chief Medical Officer of the day (J Parlane Kinloch), was no less than 'improving the quality of the Scottish race'.

The speed with which the Lanarkshire Schools Milk Experiment was planned, implemented by the Education Authority, and then reported, indicates clearly an urgency and organization rare to us today: all the field work involving 20 000 children, analysis and publication was achieved in less than 12 months. The experiment was a three-arm trial comparing the effects of  $\frac{3}{4}$  pint grade A tuberculin raw whole milk (5000 children),  $\frac{3}{4}$  pint grade A tuberculin-tested pasteurized whole milk (5000 children), and non-supplemented controls (10 000 children). The trial was designed to compare the effects of the two types of whole milk on children aged 5–12 years, located in 67 schools, in industrial and densely populated parts of Lanarkshire. Approximately one-third of the children were said to come from families in which both parents were unemployed or only partially employed. Each school

provided between 200 and 400 pupils for the experiment, half of whom were controls. The other half were supplemented with either raw milk or pasteurized milk. No school provided children covering all three arms.

Details of the selection criteria adopted for schools is not provided in the report of the study.<sup>10</sup> Selection of children within schools, although ultimately the responsibility of the school medical officers, was actually implemented by head teachers, but the section of the report about this process lacks clarity. The advice given to head teachers was that '... the selected children should be *representative of the group* (my italics) and not the most ill nourished or of any other outstanding character', and that '... controls should also be *representative of the average child* (my italics)'. It is of interest that, even with such large numbers, the importance of seeking generalizations valid across the full range of physical types of children in society played second fiddle to the perceived importance of focussing on the *average child*. This orientation towards excluding 'un-average' or 'outlier' children led to a confused selection process as 'in certain cases they selected them by ballot and in others on an alphabetical system'. Details are missing from the report, but the process was hardly helped by the (perhaps proud) admission that '... in any particular school where there was any group to which these methods had given an undue proportion of well-fed or ill-nourished children, others were substituted in order to attain a more level selection'. The comfort this engendered is well represented by the observation that '... the school medical officers were definitely of opinion that each school had furnished a very fair *average* (my italics) of its ordinary scholars ...'.

The same staff using the same instruments measured and weighed the children in their indoor clothes and without shoes before the experiment started (February) and again when it ended (June). Analysis of weight and height changes was undertaken across all schools after stratifying the 17 159 children for whom there were complete data into year age groups from 5 to 11 (there were insufficient numbers of children aged 12). After acknowledging that controls were initially taller and heavier than those in the two supplemented groups, correlations were undertaken in an attempt to demonstrate that initial size had little impact on subsequent weight and height changes. The results were interpreted as demonstrating that, for both these indices, the milk supplemented groups were superior to controls; that neither age nor gender confounded this finding; and that the effects of raw and pasteurised milk were '... so far as we can judge, equal'.

The logistics of delivering two types of milk simultaneously to so many schools each day presumably played a part in the decision to run only two of the three arms in each school. It was not long before no less an

authority than the statistician Ronald Fisher raised this issue as a design fault, especially in regard to the conclusion that raw and pasteurised milk had had similar effects.<sup>11,12</sup> On closer inspection, raw milk appeared to have a greater effect, raising the possibility that at least part of the 'milk effect' could be caused by substances other than fat, protein or sugar (which do not differ between the two milk types).

Later the same year, a beautifully written statistical critique was published by WS Gosset, under his famous pseudonym of 'Student', systematically examining the design faults in the study and their implications.<sup>13</sup> These problems included not only the issue of two rather than three arms per school, but also: selection bias resulting from the quasi-random group allocation procedures; measurement bias arising from probable variation between groups in clothing; the failure to match analysis across schools; and the erroneous conclusions about the purported lack of differential of raw and pasteurized milk. Student conceded that there probably was an effect of milk on height and weight, whilst noting that '... the conclusion ... shifted from the sure ground of scientific inference to the less satisfactory foundation of mere authority and guesswork by the fact that the "controls" and "feeders" were not randomly selected'. It seems probable that the intervention groups were, with the best of intentions, biased in favour of the smaller, thinner or less well-nourished schoolchildren, thereby biasing the comparison against the milk supplements, as well as confounding the generalizability of the results to the whole population. Student concluded that a smaller, more carefully designed study using identical twins would have been preferable.

I encourage readers who are interested in the impact of these community trials on subsequent school meal and school milk policies in the UK, and the debate about whether this was a welfare or an educational matter (or of

industrial economic significance), to consult reports of the research on these aspects which has been reported by Peter Atkins at the University of Durham.<sup>2,3</sup>

*Acknowledgments* I am most grateful for the expert writings, advice and assistance of Peter Atkins, University of Durham.

## REFERENCES

- 1 Petty EC. *The Impact of The Newer Knowledge of Nutrition: Nutrition Science and Nutrition Policy, 1900–1939*. PhD Thesis (DX188569). London: London University, 1987
- 2 Atkins PJ. Fattening children or fattening farmers? School milk in Britain, 1921–1941. *Econ Hist Rev LVIII* 2005;I:57–78
- 3 Atkins PJ. The Milk in Schools Scheme, 1934–45: 'nationalization' and resistance. *Hist Educ* 2005;34:1–21
- 4 McCollum EV. The nutritional value of milk. In: Rogers LA, Lenoir RD, eds. *World's Dairy Congress, Washington DC, 2–10 October 1923*. Washington: US Government Printing Office, 1924:421–37
- 5 Auden GA. An experiment in the nutritive value of an extra milk ration. *J R Sanitary Inst* 1923;44:236–47
- 6 Corry Mann HC. *Diets For Boys During The School Age*. Medical Research Council Special Report Series No. 105. London: HMSO, 1926
- 7 Orr JB. Milk consumption and the growth of school-children. *Lancet* 1928;i:202–3
- 8 Leighton G, Clark M. Milk consumption and the growth of school-children. *Lancet* 1929;i:40–3
- 9 Orr JB, Leighton G. The Scottish milk-feeding investigation in schools. *J State Med* 1929;37:524–7
- 10 Leighton G, McKinlay P. *Milk Consumption and The Growth of School Children*. Edinburgh: Department of Health for Scotland and London: HM Stationery Office, 1930
- 11 Fisher RA, Bartlett S. Pasteurised and raw milk. *Nature* 1931;127:591–2
- 12 Bartlett S. Nutritional value of raw and pasteurised milk. *J Ministry Agric* 1931;38:60–4
- 13 Student. The Lanarkshire Milk Experiment. *Biometrika* 1931;23:398–406